Structural Characterization of Triple Transmembrane Domain Containing Fragments of a Yeast G Protein-Coupled Receptor in an Organic: Aqueous Environment by Solution-State NMR Spectroscopy

Fracchiolla, Katrina ; Cohen, Leah ; Arshava, Boris ; Poms, Martin ; Zerbe, Oliver ; Becker, Jeffrey M ; Naider, Fred

Abstract: This report summarizes recent biophysical and protein expression experiments on polypeptides containing the N-terminus, the first, second and third transmembrane domains and the contiguous loops of the -factor receptor Ste2p, a G protein-coupled receptor. The 131-residue polypeptide Ste2p(G31-R161), TM1-TM3 was investigated by solution NMR in trifluorethanol/water: TM1-TM3 contains helical transmembrane domains at the predicted locations, supported by continuous sets of medium-range NOEs. In addition, a short helix N-terminal to TM1 was detected, as well as a short helical stretch in the first extracellular loop. Two 161-residue polypeptides, [Ste2p(M1-R161), NT-TM1-TM3], that contain the entire N-terminal sequence, one with a single mutation, were directly expressed and isolated from E. coli in yields as high as 30 mg/L. Based on its increased stability, the L11P mutant will be used in future experiments to determine long-range interactions. The study demonstrated that 3-TM domains of a yeast GPCR can be produced in isotopically labeled form suitable for solution NMR studies. The quality of spectra is superior to data recorded in micelles and allows more rapid data analysis. No tertiary contacts have been determined, and if present, they are likely transient. This observation supports earlier studies by us that secondary structure was retained in smaller fragments, both in organic solvents and in detergent micelles, but that stable tertiary contacts may only be present when the protein is imbedded in lipids.

DOI: https://doi.org/10.1002/psc.2750

Structural Characterization of Triple Transmembrane Domain Containing Fragments of a Yeast G Protein-Coupled Receptor in an Organic:Aqueous Environment by Solution-State NMR Spectroscopy

Journal:	*Journal of Peptide Science*
Manuscript ID:	PSC-14-0231.R1
Wiley - Manuscript type:	Special Issue Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Fracchiolla, Katrina; College of Staten Island, CIty University of New York, Chemistry Department; The Graduate Center, City University of New York, Department of Biochemistry Cohen, Leah; College of Staten Island, City University of New York, Department of Chemistry Arshava, Boris; College of Staten Island, CIty University of New York, Chemistry Department Poms, Martin; University of Zurich, Institute of Organic Chemistry Zerbe, Oliver; University of Zurich, Institute of Organic Chemistry Becker, Jeffrey; Univeristy of Tennessee, Department of Microbiology Naider, Fred; College of Staten Island, City University of New York, Chemistry Department; The Graduate Center, City University of New York, Department of Biochemistry
Keywords:	GPCR fragments, biosynthesis, isotopic labeling , solution-state NMR

Structural Characterization of Triple Transmembrane Domain Containing Fragments of a Yeast G Protein-Coupled Receptor in an Organic: Aqueous Environment by Solution-State NMR Spectroscopy

Katrina E. Fracchiolla[1,2], Leah S. Cohen[1], Boris Arshava[1], Martin Poms[3], Oliver Zerbe[3], Jeffrey M. Becker[4], Fred Naider[1*]

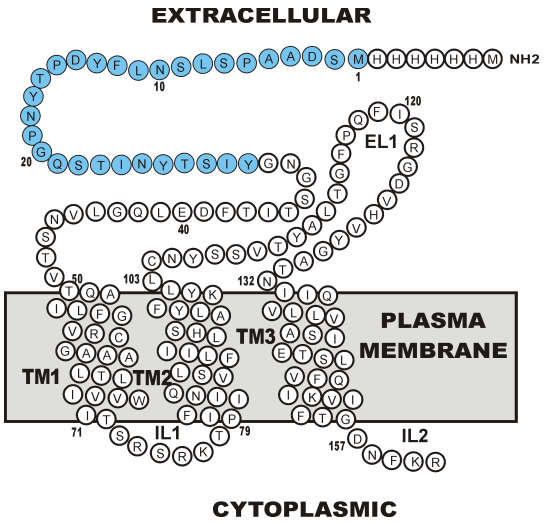
[1]Department of Chemistry, The College of Staten Island, City University of New York (CUNY), Staten Island, NY 10314; [2]Department of Biochemistry, The Graduate Center, CUNY, New York, NY, 10016; [3]Institute of Organic Chemistry, University of Zurich, Switzerland; [4]Department of Microbiology, University of Tennessee, Knoxville, TN 37996

Short Running title: Characterization of Triple TM Domains of a GPCR

*) Department of Chemistry, The College of Staten Island, CUNY, 2800 Victory Boulevard, Staten Island, NY 10314, USA. Tel.: +1 718 982 3896; fax: +1 718 982 3910. E-mail address: fred.naider@csi.cuny.edu (F. Naider). Fred Naider is the Leonard and Esther Kurtz Term Professor at the College of Staten Island, City University of New York

**) Invited Article for the Anniversary Issue of Journal of Peptide Science. It is dedicated to Murray Goodman in memory of his strong support for this journal as member of the editorial board (1995-??)

Graphical Abstract



Structural characterization of integral membrane proteins is an arduous task that requires patience and innovation for progress. In this paper we characterize the secondary structure of a three transmembrane domain fragment of the Ste2p G protein-coupled receptor from *Saccharomyces cerevisiae*. Methods are introduced to increase expression and stability of this protein for future structural characterization.

Abstract

This report summarizes recent biophysical and protein expression experiments on polypeptides containing the N-terminus, the first, second and third transmembrane domains and the contiguous loops of the α-factor receptor Ste2p, a G protein-coupled receptor. The 131-residue polypeptide Ste2p(G31-R161), TM1-TM3 was investigated by solution NMR in trifluorethanol/water: TM1-TM3 contains helical transmembrane domains at the predicted locations, supported by continuous sets of medium-range NOEs. In addition, a short helix N-terminal to TM1 was detected, as well as a short helical stretch in the first extracellular loop. Two 161-residue polypeptides, [Ste2p(M1-R161), NT-TM1-TM3], that contain the entire N-terminal sequence, one with a single mutation, were directly expressed and isolated from *E. coli* in yields as high as 30 mg/L. Based on its increased stability, the L11P mutant will be used in future experiments to determine long-range interactions. The study demonstrated that 3-TM domains of a yeast GPCR can be produced in isotopically labeled form suitable for solution NMR studies. The quality of spectra is superior to data recorded in micelles and allows more rapid data analysis. No tertiary contacts have been determined, and if present, they are likely transient. This observation supports earlier studies by us that secondary structure was retained in smaller fragments, both in organic solvents and in detergent micelles, but that stable tertiary contacts may only be present when the protein is imbedded in lipids.

Key words: GPCR fragments, biosynthesis, isotopic labeling , solution-state NMR

Abbreviation List: G protein-coupled receptor, GPCR; transmembrane, TM; nuclear magnetic resonance, NMR; Ste2p(G31-T110), TM1-TM3; trifluoroethanol, TFE; Ste2p (M1-R161), NT-TM1-TM3; Luria Broth, LB; ampicillin, Amp; isopropyl β-D-1-thiogalactopyranoside, IPTG; reverse phase high performance liquid chromatography, RP-HPLC; Circular Dichroism, CD; sodium dodecyl sulfate, SDS; lysomyristoylphosphatidylglycerol, LMPG; N-terminus, NT; extracellular loop 1, EL1; paramagnetic relaxation enhancements, PREs; Ste2p (M1-R161, L11P), NT^{L11P}-TM1-TM3.

Introduction

In the early 1960s the laboratory of Professor Murray Goodman initiated a seminal series of studies using synthetic homo-oligopeptides to understand aspects of peptide and protein structure. These studies led to fundamental insights into the chain length dependence of the formation of α-helices [1]. The initial studies were significantly expanded by the Toniolo group to provide information on β-structure formation both in solution and in the solid state [2, 3]. Most importantly the use of synthetic peptides as surrogates to probe questions related to protein structure was established. Since these early investigations thousands of investigations have been conducted on the conformational preferences of carefully designed peptides leading to breakthroughs in our understanding of β-hairpin or β-turn formation and the assumption of structure by β- and γ-peptides [4-7].

In this report we present studies on folding of a GPCR using a 3-transmembrane containing peptide surrogate of the α-factor receptor from the yeast *Saccharomyces cerevisiae*. This report describes our most recent work on Ste2p peptide surrogates, which began in the early 1990s, some of which was recently reviewed [8]. Here we report conformational preferences of a 130-residue peptide corresponding to the first 3 TM domains of the Ste2p receptor and two of its loop regions encompassing residues 31-161 in an organic:aqueous medium that was used as a membrane mimic previously [9]. This work reflects an outgrowth of training and an approach that one of us (FN) received in the Goodman laboratory and it is dedicated to the memory of Dr. Goodman and to the myriad of scientists who were trained under his mentorship.

Structural characterization of G protein-coupled receptors is notoriously difficult due to the inherent properties of these receptors, with high-resolution crystallographic structural information available to date for slightly more than 20 of the ~1000 identified GPCRs [10-40]. Modifications of most of these proteins through introduction of conformation-stabilizing mutations resulting in higher melting temperatures, crystallization in presence of antibodies, truncation of flexible segments, and insertion of a crystallization-nucleation protein into the long and flexible third intracellular loop were necessary in order to facilitate stability and crystal packing interactions. Although crystallographic studies of GPCRs have provided fundamental information, most of those

modifications compromise the activity of the proteins, and therefore sparse information on the dynamics of the protein has been obtained so far. Accordingly, nuclear magnetic resonance (NMR) spectroscopy investigations are a valuable complement to crystallographic analyses.

NMR investigations of GPCRs are hampered by a number of technical issues associated with the production of the large quantities of isotopically labeled receptor required for NMR in the expression hosts that were used by crystallographers, by the difficulty of establishing conditions that mimic the biologically relevant environment of the receptor while still providing good-quality spectra, by the tendency of these membrane proteins to aggregate, and by the large size of the protein/lipid complexes. To date, NMR investigations have been reported for heptahelical integral membrane proteins, the vasopressin receptor, the CXCR1 receptor, the CB2 receptor, the β2-adrenergic receptor, and sensory rhodopsin and proteorhodopsin [41-48]. The studies on the CB2 and β2-adrenergic receptors have focused on dynamics utilizing a mixture of NMR and molecular stimulations to better understand conformational changes [47, 48]. High-resolution solution state structures have been reported for the bacterial GPCR analogues sensory rhodopsin and proteorhodopsin [46, 49]. These proteins are somewhat smaller in size when compared to most GPCRs, less flexible, and hence more well-behaved. Solid-state NMR was used for the determination of the NMR structure of the mammalian CXCR1 receptor in phospholipid bilayers[43]. While solid state approaches hold promise for the future, an NMR structure for a full-length, mammalian GPCR based strictly on measured constraints has yet to be determined.

In an effort to overcome some of the difficulties associated with the NMR characterization of full-length GPCRs, several groups have focused on the characterization of *fragments* of GPCRs. Fragments are often easier to express in high yields, and the smaller number of residues leads to less crowded spectra. Our group studies the yeast α-factor receptor, Ste2p, a 431-residue peptide ligand receptor, which we are using as a model system for GPCR methods development. We have published the only solution structure for a GPCR fragment containing two TMs [TM1-TM2; Ste2p(G31-T110)] in LPPG micelles and in 2,2,2-trifluoroethanol (TFE):water mixtures [9, 50]. In both cases, the fragment is helical and forms a hairpin. However, the helical

hairpin is more stable in LPPG and only transiently formed in TFE:water. The formation of a tertiary structure, even a transient tertiary structure, supports the hypothesis that large domains of a GPCR can fold independently of the remainder of the protein.

All X-ray structures of GPCRs show that every TM domain is in contact with at least two other TM domains. Therefore, we hypothesized that increasing the size of our Ste2p fragment to 3TM domains would increase the probability of forming tertiary contacts and potentially result in a more stable structure through increased mutual stabilization. As a result, we expanded our structural characterization to a 3TM containing fragment of Ste2p(G31-R161), TM1-TM3. This fragment contains 131 residues of Ste2p, including 19 residues from the N-terminal domain, the first TM through the third TM with connecting loops and five residues of the second intracellular loop. Here we report details of a structure and dynamics study on Ste2p TM1-TM3 in 50% TFE:water. Recently, we showed that the addition of the first 30-residues of the Ste2p N-terminus increased expression and the stability of Ste2p TM1-TM2 in NMR preparations [8]. We will also report on the expression and biophysical characteristics of Ste2p (M1-R161) NT-TM1-TM3, which contains 161-residues of Ste2p including the entire N-terminal domain and the same TMs and loops as above.

Materials and Methods

Assignment of Side Chain Resonances

NMR backbone assignment of the TM1-TM3 fragment of Ste2p in TFE:water at 45°C was previously reported [51]. Side chain resonances were assigned using the HCCH-TOCSY [52, 53], HCCC(CO)NH [54], and (HM)CM(CGCBCA)NH and (HM)CM(CBCA)NH [55] experiments using NMRView 5 [56] and CARA [57]. Briefly, Cα and Cβ annotations from the backbone assignments were confirmed in the HCCC(CO)NH spectra. The latter were also useful to obtain frequencies of the connected protons. Sidechain assignments of aliphatic resonances were then completed with the help of HCCH-TOCSY spectra starting from anchoring resonances in the 2D [^{13}C,^{1}H]-HSQC experiments. In general, the [^{13}C,^{1}H]-HSQC spectrum was very crowded, and assignment of sidechain resonances using the CA and CB chemical shifts was difficult. Assignments of methyl groups in the ILV-labelled sample was performed using experiments published by the Kay group [55, 58] that start on methyl protons and connect to amide moieties. Knowledge of methyl assignments then facilitated sidechain assignments via HCCH-TOCSY correlations form the methyl moieties. The spectra were acquired using either a three-channel Varian NMR-S 600 MHz NMR spectrometer (Varian NMR Instrument, Palo Alto, CA) with a z-axis pulsed-field-gradient and a Varian 5mm [^{1}H, ^{15}N, ^{13}C, ^{2}D] cryo-probe at the College of Staten Island, a three-channel Bruker AV-700 700 MHz NMR spectrometer (Bruker, Billerica, MA) equipped with a CRYO TXI inverse triple resonance cryoprobe at the University of Zurich, or a four-channel Bruker 800 MHz NMR spectrometer (Bruker, Billerica, MA) equipped with a CRYO TCI triple resonance cryoprobe at the New York Structural Biology Center.

Confirmation of Secondary Structure Localization in Ste2p TM1-TM3 (G31-R161) using ^{15}N T_2 relaxation and H,D Amide Exchange

^{15}N T_2 relaxation experiments were performed on a 0.5 mM [^{15}N]-TM1-TM3 sample solubilized in 50% TFE-d_2:(water+0.1% TFA) (Sigma, St. Louis, MO). A series of eight [^{15}N,^{1}H]-HSQC-based CPMG experiments were performed with varying relaxation times of 0, 10, 30, 50, 70, 110, 150 and 210 ms. The data were processed and

the rate analysis function of NMRView5 was used to calculate the relaxation time for each residue.

H,D amide exchange analysis was performed on a 0.5 mM [^{15}N]-TM1-TM3 sample obtained by dissolving lyophilized protein in a fully deuterated solvent 50% TFE-d3:(D_2O + 0.1% TFA-d1) (Sigma, St. Louis, MO). A series of [^{15}N,^{1}H]-HSQC experiments were measured at 40 minute intervals for a total of 5 hours. Additional spectra were collected daily with the final spectrum collected eight days after the original sample preparation. The rate analysis function of NMRView5 was used to calculate the exchange time for each assigned residue, resulting in a logarithmic plot of the exchange time vs. residue number (see Figure 1A).

Assessment of the Relative Flexibility of Ste2p TM1-TM3 (G31-R161) using an Analysis of the $^{15}N\{^{1}H\}$-NOE

A 0.5 mM sample of [^{15}N]-TM1-TM3 was solubilized in 50% TFE-d_2:water and subject to analysis of the $^{15}N\{^{1}H\}$-NOE both at 45°C and 30°C. Amplitudes and volumes for each crosspeak in each data set were calculated using CARA. The $^{15}N\{^{1}H\}$-NOE was derived by computing the ratio of peaks in the spectra with and without prior proton irradiation.

NOESY Assignment and Structure Calculation for Ste2p TM1-TM3(G31-R161)

Backbone chemical shift assignments were used to calculate torsion angles for TM1-TM3 using TALOS+[59]. The ATNOS-CANDID[60, 61] component of the UNIO suite, that is interfaced with the structure calculation program CYANA[62], was used for automated assignment of the 3D [^{13}C]- and [^{15}N]-resolved NOESY spectra based on the assigned backbone and sidechain chemical shifts. A seven-cycle CYANA iteration using the 3D [^{15}N]- and [^{13}C]-resolved NOESY spectra as well as the TALOS-derived torsion angle restraints was used to calculate 80 structures, and the 20 lowest energy structures for TM1-TM3 were analyzed in detail.

Direct Expression of NT-TM1-TM3 Protein Fragments

A construct containing an N-terminal His_6-tag, followed by the full N-terminal tail and an S104C mutation (NT-TM1-TM3, Ste2p(M1-R161, S104C) corresponding to Ste2p(M1-R161) was cloned and expressed. The plasmid containing the NT-TM1-TM3 sequence was transformed into BL21(DE3), 20 μL of the transformation reaction was used to inoculate 50 mL of Luria Broth containing ampicillin (LB/Amp). The culture was grown overnight at 37°C and used to inoculate 1 L of LB/Amp, and the cells were grown at 37°C until an OD_{600} of approximately 1 was reached. The cells were then pelleted at 4000 rpm for 20 minutes and resuspended in M9 minimal medium using ^{13}C-glucose and $^{15}NH_4Cl$ as the sole carbon and nitrogen sources as needed, respectively (Cambridge Isotope Laboratories, Andover, MA). Expression was induced by the addition of 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and the culture was incubated at 30°C for 22h, after which cells were pelleted at 5000 rpm for 20 min. Inclusion bodies were prepared from the cell pellets as previously described [63], with the addition of 5 mM dithiothreitol to the lysis buffer. Inclusion body pellets were dissolved in 70% TFA by sonication and purified by preparative reverse phase high performance liquid chromatography (RP-HPLC) on a Zorbax 300SB-C3 Prep HT (Agilent, Santa Clara, CA) 21.2 x 150 mm, 7 micron column at 60°C using gradient elution with an acetonitrile/2-propanol gradient from 30% to 72% Solvent B (80% acetonitrile, 20% 2-propanol, 0.1% TFA) where Solvent A contained 80% water, 20% 2-propanol, 0.1% TFA). After purification, ^{15}N, $^{15}N^{13}$C and $^{15}N^{13}C^{2}$H-labelled proteins were lyophilized.

Circular Dichroism Spectroscopy of Ste2p TM1–TM3 and NT-TM1-TM3 Peptides.

CD spectroscopy was used to access the secondary structure of the TM1–TM3 peptides. Experiments were carried out in both organic:aqueous solvents and detergents. Peptides were solubilized by sonication in 50% TFE:(water+0.1% TFA) and the concentration of the peptide stock solution was determined by UV absorbance at 280nm. The molar extinction coefficient of 12,950 $M^{-1}cm^{-1}$ was used for short M1M3 peptide and 19035 $M^{-1}cm^{-1}$ for NT-TM1-TM3 and NT^{L11P}-TM1-TM3. The extinction coefficients were calculated using 1490 for Tyr, 5500 for Trp and 125 for Cys residues. This stock solution was portioned and lyophilized to yield approximately 20 μM or 7 μM solutions in detergent and TFE:water media, respectively. The detergents were prepared as 20 mM

sodium dodecyl sulfate (SDS) (Sigma, St. Louis, MO) or lysomyristoylphosphatidylglycerol (LMPG) (Avanti Polar Lipids, Alabaster, Alabama) in 20 mM sodium phosphate buffer, pH 5.0. All samples were sonicated at 40 ^{o}C for 5 min at 40 W using a S3000 sonicator (Farminigdale, NY) with a cup horn. The spectra were recorded on an Aviv Model 410 CD instrument (Aviv Biomedical, Lakewood, NJ). For the organic:aqueous studies, a quartz cuvette with a pathlength of 1 mm was used. For the detergent studies, a quartz cuvette with a pathlength of 0.2 mm was used. The spectra were collected over a wavelength range of 260 nm to 190 nm in increments of 1 nm. An average of four scans was taken in all cases, and the background subtraction used the spectrum from the solvent medium. The raw data were then converted in mean residue molar ellipticity (deg cm^2 dmol $^{-1}$). Deconvolution analysis was performed using the CDNN software [64].

NMR Analysis to Determine Peptide Stability at High Concentrations.

For NMR investigations in organic:aqueous media, 2-3 mg of lyophilized protein were solubilized in 175 µL of TFE and sonicated as above. After sonication, 175 µL of water containing 0.1% TFA were added and the sample was sonicated again. A clear solution was obtained and the sample was transferred to a Shigemi NMR tube (Shigemi, Allison Park, PA). ^{1}H NMR experiments were conducted at 45 °C on a three-channel Varian NMR-S 600 MHz NMR spectrometer (Varian NMR Instrument, Palo Alto, CA) with a z-axis pulsed-field-gradient and a Varian 5-mm [^{1}H, ^{15}N, ^{13}C, ^{2}D] cryo-probe. Sample stability was assessed by measuring the overall peak integral for amide NH, NH_2 and aromatic area from 6.0 to 9.5 ppm for samples incubated for several days at 45°C.

Results and Discussion

Assignment of Side Chain Resonances of TM1-TM3

Side chain assignments were conducted using the HCCH-TOCSY, HCCC(CO)NH, (HM)CM(CGCBCA)NH, and (HM)CM(CBCA)NH experiments as described above. Data from all three experiments were combined to assign all [^{13}C,^{1}H]-HSQC crosspeaks. Complete side chain assignments were obtained except for the aromatic residues. A table of all chemical shift assignments is provided in the Supplemental Materials (Supplemental Material Table S1) and chemical shifts have been added to the BMRB database under accession code 17211.

Assessment of the Secondary Structure of Ste2p TM1-TM3(G31-R161) by H,D Amide Exchange and ^{15}N T_2 Relaxation Experiments

H,D amide exchange and ^{15}N T_2 relaxation experiments were conducted in order to gain additional insight into the stability of secondary structure of the TM1-TM3 fragment in 50% TFE:water. Analysis of the H,D exchange results (Figure 1A) reveals that, in general, the residues in the predicted TM helices (boxed regions) displayed reduced H,D exchange rates when compared to those in the loop regions. The majority of the exchange times in the predicted helices range from approximately 10-200 h while they are between 0.02-1 h in the loop regions. Accelerated exchange is also observed in the middle of TM1, centered around the GXXXG motif. The predicted TM helix boundaries[51] seem to correlate very well with the exchange data for residues of TM1 and TM3, as slow exchange is only observed within the TM regions (indicated by boxes in Figure 1A), while fast exchange is restricted to loop regions. Except for residues in the very center of TM2, the exchange data for TM2 reveal that many residues near the ends of the helix possess exchange times between 1 and 10 h indicating reduced stability. This finding deviates from the conclusions previously reported that were derived from chemical shift analysis[51]. The putative TM3 displays more stable hydrogen bonds at its N-terminus, however the differences in exchange times across the helix are, in general, not as large as in TM2. In addition to the TM helices, the chemical shift analysis indicated the presence of helices encompassing residues 38 to 49 or 108 to 115 within the N-terminus (NT) and the extracellular loop 1 (EL1), respectively[51]. These helices are

not detected in the H,D exchange data indicating that they are rather unstable.

^{15}N T_2 relaxation times (Figure 1B) are between 25 and 50 ms in the TM segments, and adopt values on the order of hundreds of milliseconds at the termini and around 50 ms in the loops. We observed increased relaxation times for the GXXXG region in TM1. TM1 again appears to be N-terminally extended, and this extension correlates well with the position of the N-terminal helix. Interestingly, relaxation times for residues in EL1 gradually increase until residue 120. This gradual increase is consistent with the presence of a short helix in the beginning of EL1. Furthermore, the relaxation times in TM2 are the shortest of the 3 TM regions. Overall, the relaxation results seem to confirm the helix boundaries previously identified from secondary chemical shifts [51]. We note that exchange data allows us to differentiate different degrees of helix stability for residues that may display very similar relaxation data as the latter in addition to helix stability may contain contributions from conformational exchange.

Assessment of the Relative Flexibility of Ste2p TM1-TM3(G31-R161) by Analysis of the $^{15}N\{^1H\}$-NOE

Analysis of the $^{15}N\{^1H\}$-NOE has also been used to gain insights into the flexibility of membrane proteins [65]. Based on previous studies on the TM1-TM2 fragment of Ste2p [9] it was expected that the TM1-TM3 construct would be more structured at lower temperature. Accordingly, a series of HSQC experiments were run in 5°C decrements from 40°C to 25°C in order to transfer the backbone amide assignments (data not shown). It was found that 30°C was the lowest temperature at which chemical shift adaptations could be performed. Accordingly, $^{15}N\{^1H\}$-NOE experiments were performed at both 45°C and 30°C.

$^{15}N\{^1H\}$-NOE data collected at 45°C (Figure 2, top) closely reflect trends in the T_2 data. Both the N and C-termini display negative or small positive values, suggesting that these regions of the fragment are highly flexible. The three major regions with $^{15}N\{^1H\}$-NOE above 0.5 fall within the putative TM regions (dashed lines, Figure 2), but also include the short helix N-terminal to TM1. There is a small dip at the GXXXG motif in TM1. The putative loop regions show increased mobility relative to the TM

regions. Again, the N-terminus of EL1 displays decreased mobility in agreement with the presence of some degree of order in that part of this loop.

In general, data at 45°C seem to contain less noise in comparison to the 30°C data, likely due to the fact that peaks can be integrated more reliably at the higher temperature at which the lines are sharper and spectral overlap is less of a problem. At 30°C, the $^{15}N\{^{1}H\}$-NOE for most regions of TM1-TM3 increases slightly. For example, TM1 appears to become less mobile at decreased temperature (many ratios >0.7), while the trend for increased mobility within the GXXXG motif is retained. We believe that, in general, the data supports increased overall rigidity for the protein at 30°C. The most significant differences are observed for EL1, for which the C-terminal half becomes rigidified as reflected in an increase of the $^{15}N\{^{1}H\}$-NOE from 0 to approximately 0.25. We attribute the effect to a stabilization of secondary structure triggered by formation of tertiary contacts. The more frequent formation of these contacts results in part in exchange broadening of resonances, resulting in the observed problems in integration of spectra and deterioration of spectra quality at the lower temperature.

NOESY Assignment and Structure Calculation for Ste2p TM1-TM3(G31-R161)

A total of 342 intraresidual ($|i-j|=0$), 531 sequential ($|i-j|=1$), 873 short-range ($|i-j| \leq 1$), 544 medium-range ($1<|i-j|<5$), and 5 long-range ($|i-j|\geq5$) restraints were used in the calculation (Supplemental Material Figure S1 and Table S2). The twenty lowest-energy of the computed 80 structures were used to evaluate the structure. They had target function values ranging from 0.6 to 1.41. A total of 10 violations were observed (9 distance and 1 van der Waals violations). 90.3 % of dihedrals fell within the most favored regions of the Ramachandran plot, with the remaining 9.7% located in the additionally allowed regions (Supplemental Material Figure S2). As expected, the majority of the residues in the most favored region were in the right-handed α-helix portion of the plot.

An analysis of typical NOE contacts identified revealed a large number of *i, i+3* contacts throughout the N-terminus-TM1 region (Figure 3) for residues 36-49, 51-59, and 63-74. The absence of *i, i+4* contacts for residues 49-51, that corresponds to the junction between the N-terminal helix and TM1, suggests a destabilization of the α-helix at this

position. Similarly, the absence of *i, i+4* contacts around the GXXXG motif in TM1 is consistent with reduced secondary chemical shifts, accelerated H,D exchange, and decreased T_2 relaxation.

The boundaries of the second helix are consistent with a C-terminal extension. This extension may also be the result of the helix-inducing properties of TFE:water, as it most likely represents an elongation of TM2 by part of EL1. Absence of *i, i+4* contacts for residues 102-105 indicates the transition from TM2 to EL1. Finally, the boundaries of TM3 are consistent with conclusions from sequence-based secondary structure prediction software [66-69].

As very few long-range contacts were identified in the NOESY data, the final calculated conformers showed little convergence resulting in a large RMSD when superimposing the entire sequence. Therefore we superimposed only single TMs or even parts of them in the following analysis (Figure 4A, Table 1). The helices spanning the NT and TM1 (NT-TM1) and TM3 superimpose with backbone RMSDs of 3.02 for and 0.99 Å, respectively. As expected RMSDs improved when excluding segments for which no *i, i+4* contacts were observed (Figure 4B, Table 1). The resulting backbone RMSDs were 2.36 ± 1.16 Å for TM1, 1.03 ± 0.41 Å for TM2, 1.49 ± 0.51 Å for TM3, and 0.43 ± 0.18 Å and 0.69 ± 0.28 Å for the short NT and EL1 helices, respectively. The higher RMSD in TM1 is likely due to fluctuations about the central GXXXG motif. RMSD calculations performed using the TM boundaries from the sequence-based prediction software (Figure 4C, Table 1) and those derived from the homology model from Eilers [70] (Figure 4D, Table 1) resulted in the smallest RMSDs but these superimposed segments are also the shortest. Unfortunately, significant peak broadening at 25°C, did not allow recording of useful spectra under these conditions. This precluded the detection of NOEs that would be expected if tertiary structure would be present at this temperature. The secondary structure determined based on NOEs is more accurately reflected in the T_2 data in comparison to H,D amide exchange.

Our previous work has indicated that transient tertiary structure can be formed at lower temperatures (25°C) in an organic:aqueous solvent [9] but that more stable tertiary structures are formed in micellar environments [50]. Whether this also the case for TM1-TM3 is presently under investigation in our laboratories.

Direct Expression and Purification of NT-TM1-TM3: Maximizing Expression through Protein Engineering

Despite extensive attempts to optimize NMR conditions and obtain reproducibly spin-labeled 131-residue TM1-TM3 for measuring paramagnetic relaxation enhancements (PREs), the expression of the TM1-TM3(S104C) mutant proved extremely challenging. Direct expression led to very low and unreproducible yields, and CNBr cleavage of a TrpΔE fusion S104C protein resulted in no recoverable product. In contrast, following a recently developed approach for TM1-TM2 [8], direct expression of the longer Ste2p(Met1-R161, S104C) [NT-TM1-TM3] without fusion to a leader sequence (Supplementary Material Figure 3A) doubled yields to up to 30 mg per liter of culture. Expression was not always reproducible, however, and sometimes we experienced trouble with NMR sample preparation where the protein would go into solution but the solution would become gelatinous or the protein would precipitate over time. An additional mutation in the N-terminus, L11P (NT^{L11P}-TM1-TM3), resulted in more reproducible expression yields and stable protein samples (*vide infra*) (Supplementary Material 3B). Previous work demonstrated that mutations in the N-terminus of Ste2p may result in changes in surface expression and signaling of this receptor *in vivo* [71-73]. SCAM analysis has indicated that downstream of residue 11 there is a β-sheet present [72, 73] and this region was also predicted to have β-strand potential between L8 and T16 [74-76]. The presence of proline at position 11 can disrupt this conformation. Furthermore, recent work in the Becker lab on the functional implications of NT deletions revealed that Ste2p constructs missing residues 1-10 or 11-20 have increased biological function as indicated by reporter gene activity and/or mating efficiency (unpublished results). Since the N-terminus has been implicated in Ste2p dimerization [73] the destabilization of the β-strand by the Leu11Pro could have resulted in decreased aggregation resulting in more reliable expression yields and protein samples that will be amenable to NMR analysis. Our preliminary results demonstrate that this is the case for NT^{L11P}-TM1-TM3.

Biophysical characterization of NT-TM1-TM3 and comparison to TM1-TM3

CD spectroscopy was used to determine if the presence of the NT changes the overall secondary structure and fragment stability in TFE:water samples or micellar solutions. Previous analysis has indicated that a correlation between CD spectra and NMR analysis is useful to determine solvent conditions for NMR. The CD patterns of TM1-TM3, NT-TM1-TM3, and NT^{L11P}-TM1-TM3 were measured in two different TFE:water mixtures and micellar preparations. Furthermore, these samples were analyzed over a week to determine stability of the solution.

No difference in CD spectra was observed between 50% TFE:water and 67% TFE:water for the three protein samples TM1-TM3, NT-TM1-TM3, and NT^{L11P}-TM1-TM3 (Figure 5, left panels). The resulting spectra were analyzed for secondary structure tendencies using the deconvolution program CDNN[64] (Table 2). There was an overall decrease in helicity of the NT containing peptides which is likely due to an increase in the amount of β-sheet tendency in these peptides as described above. The β-strand percentage in NT^{L11P}-TM1-TM3 is decreased, presumably due to the presence of Pro11 which should lower the β-strand tendency of the NT. As a result, the overall helicity of this peptide fragment increases when compared to the same peptide without the proline mutation. Similar structural tendencies are observed in micellar environments (Figure 5, right panels, Table 2). However, the amount of β-sheet structure in micellar environments compared to the organic aqueous mixtures is higher in all of the peptides. No differences were observed after one week at room temperature indicating that at the low concentration used in CD measurements, the protein samples are stable.

Preliminary investigations of NMR sample conditions have shown that the NT-TM1-TM3 peptides are more stable in 67% TFE:water than in 50% TFE:water. Under similar conditions and concentrations the sample of NT^{L11P}-TM1-TM3 remains clear for longer time than TM1-TM3 or NT-TM1-TM3. After 6 days of incubation at 45°C, the best behaving sample of NT-TM1-TM3 lost ~30% of the peak intensity in both the NH/aromatics region and the aliphatic region. Conversely, the NT^{L11P}-TM1-TM3 sample retained 86% intensities in these same regions after one week. Most important is our observation that the NT^{L11P}-TM1-TM3 sample could be reproducibly prepared and remained NMR-stable for at least one week. This type of stability would allow for measurements of the 3D and 4D NMR experiments. This was not the case for NT-TM1-

TM3 lacking Pro11. Pro11 may decrease the aggregation propensity of the N-terminus leading to a sample that is more suitable for NMR analysis. This observation may have implications for the development of a full-length Ste2p construct for NMR investigations.

Conclusions

Extensive NMR characterization of the TM1-TM3 fragment of Ste2p has been conducted in 50% TFE:water. Complete backbone and sidechain assignments have been made for data collected at 45°C. The backbone assignments were used for qualitative secondary structure and dynamics analysis via chemical shift analysis, as well as for H,D amide exchange, ^{15}N T_2 relaxation, and $^{15}N\{^1H\}$-NOE measurements. All analyses indicate the presence of three TM helices, with boundaries in close agreement with sequence-based predictions, with additional helicity observed in the N-terminal region and the N-terminus of EL1. Analysis of the $^{15}N\{^1H\}$-NOE suggests that certain areas of the fragment become less mobile at reduced temperatures. Reduction in the relative mobility of EL1 at 30°C would be consistent with the conclusion that the fragment adopts some degree of tertiary structure at reduced temperature. As 30°C is the optimal growth temperature for *Saccharomyces cerevisiae*, the organism expressing Ste2p, it is tempting to speculate that these tertiary structures are biologically relevant and formed during folding of the receptor.

The NOE-restrained structure calculation using data collected at 45°C reveals the presence of 4 helices. Three of them are in good agreement with the homology and secondary chemical shift predictions. Alignment of the calculated boundaries for the TM regions and the secondary shift boundaries for the NT and EL1 helices reveals acceptable convergence and relatively low RMSD values. PRE experiments will need to be performed to identify possible long-range contacts. However, our preliminary results (data not shown) have not revealed such contacts in the organic:aqueous medium.

Our initial hypothesis was that increasing the number of TMs in a GPCR fragment would increase the probability for formation of interhelical contacts resulting in a better defined tertiary structure in TFE:water. Comparison of the TM1-TM2 and TM1-TM3 structural analyses in TFE:water suggests that additional work is necessary to test this hypothesis. TM1-TM2 adopted a transient tertiary structure with observable tertiary

contacts as judged by PRE experiments. However, to date we have not been able to run reproducible PRE experiments with the analogous TM1-TM3 construct. As the size of the fragment has increased, the degree of difficulty of the NMR analyses has also increased. To a large extent this may be due to the tendency of the larger fragments to aggregate in membrane mimetic media, even in organic:aqueous mixtures. It may also be possible that TM2 and TM3 compete for overlapping helix association sites on TM1. This is likely not the case within the entire receptor, but in the fragments part of the interhelical contacts are missing, and hence alternative association modes may occur. As we have recently suggested in review articles, the NMR analysis of GPCRs is complicated by both conformational exchange broadening and intermolecular interactions [8, 77]. Part of this conformational flexibility is related to the mode of receptor activation, and the fact that most receptors display basal activity in absence of agonists. GPCRs often have regions with β-strand potential, which would create surfaces that can lead to aggregation at the high concentrations used in NMR anlaysis. Furthermore numerous studies have concluded that GPCRs form dimers and higher protomers in membranes [as reviewed in 78, 79, 80]. Our initial work with the Leu11Pro mutant of NT-TM1-TM3 suggests that properly engineered constructs may lower the tendency to aggregate which will lead to more stable NMR sample preparations and increased spectral quality for NMR analyses.

Acknowledgements:

This work was supported by research grants GM22086 (F.N.) and GM22087 (J.M.B.) from the National Institutes of Health. Professor Naider is a member of the New York Structural Biology Center, which is a STAR center supported by the New York State Office of Science, Technology, and Academic Research.

References

1. Goodman M., Verdini A.S., Toniolo C., Phillips W.D., and Bovey F.A. Sensitive criteria for the critical size for helix formation in oligopeptides. *Proc Natl Acad Sci, U S A.***1969**; 64:444-450
2. Goodman M., Naider F., and Toniolo C. Circular dichroism studies of isoleucine oligopeptides in solution. *Biopolymers.***1971**; 10:1719-1730
3. Toniolo C. Intramolecularly hydrogen-bonded peptide conformations. *CRC Crit Rev Biochem.***1980**; 9:1-44
4. Gellman S.H. Minimal model systems for beta sheet secondary structure in proteins. *Curr Opin Chem Biol.***1998**; 2:717-725
5. Cheng R.P., Gellman S.H., and DeGrado W.F. beta-Peptides: from structure to function. *Chemical reviews.***2001**; 101:3219-3232
6. Searle M.S. Insights into stabilizing weak interactions in designed peptide beta-hairpins. *Biopolymers.***2004**; 76:185-195
7. Searle M.S., and Ciani B. Design of beta-sheet systems for understanding the thermodynamics and kinetics of protein folding. *Curr Opin Struct Biol.***2004**; 14:458-464
8. Cohen L.S., Arshava B., Kauffman S., Mathew E., Fracchiolla K.E., Ding F.X., Dumont M.E., Becker J.M., and Naider F. Guided reconstitution of membrane protein fragments. *Biopolymers.***2014**; 102:16-29
9. Cohen L.S., Arshava B., Neumoin A., Becker J.M., Guntert P., Zerbe O., and Naider F. Comparative NMR analysis of an 80-residue G protein-coupled receptor fragment in two membrane mimetic environments. *Biochim Biophys Acta.***2011**; 1808:2674-2684
10. Palczewski K., Kumasaka T., Hori T., Behnke C.A., Motoshima H., Fox B.A., Le Trong I., Teller D.C., Okada T., Stenkamp R.E., Yamamoto M., and Miyano M. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science.***2000**; 289:739-745
11. Cherezov V., Rosenbaum D.M., Hanson M.A., Rasmussen S.G., Thian F.S., Kobilka T.S., Choi H.J., Kuhn P., Weis W.I., Kobilka B.K., and Stevens R.C. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science.***2007**; 318:1258-1265
12. Rasmussen S.G., Choi H.J., Rosenbaum D.M., Kobilka T.S., Thian F.S., Edwards P.C., Burghammer M., Ratnala V.R., Sanishvili R., Fischetti R.F., Schertler G.F., Weis W.I., and Kobilka B.K. Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature.***2007**; 450:383-387
13. Jaakola V.P., Griffith M.T., Hanson M.A., Cherezov V., Chien E.Y., Lane J.R., Ijzerman A.P., and Stevens R.C. The 2.6 Angstrom Crystal Structure of a Human A2A Adenosine Receptor Bound to an Antagonist. *Science.***2008**; 322:1211-1217
14. Park J.H., Scheerer P., Hofmann K.P., Choe H.W., and Ernst O.P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature.***2008**; 454:183-187

15. Scheerer P., Park J.H., Hildebrand P.W., Kim Y.J., Krauss N., Choe H.W., Hofmann K.P., and Ernst O.P. Crystal structure of opsin in its G-protein-interacting conformation. *Nature.***2008**; 455:497-502
16. Warne T., Serrano-Vega M.J., Baker J.G., Moukhametzianov R., Edwards P.C., Henderson R., Leslie A.G., Tate C.G., and Schertler G.F. Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature.***2008**; 454:486-491
17. Chien E.Y., Liu W., Zhao Q., Katritch V., Han G.W., Hanson M.A., Shi L., Newman A.H., Javitch J.A., Cherezov V., and Stevens R.C. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science.***2010**; 330:1091-1095
18. Wu B., Chien E.Y., Mol C.D., Fenalti G., Liu W., Katritch V., Abagyan R., Brooun A., Wells P., Bi F.C., Hamel D.J., Kuhn P., Handel T.M., Cherezov V., and Stevens R.C. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science.***2010**; 330:1066-1071
19. Dore A.S., Robertson N., Errey J.C., Ng I., Hollenstein K., Tehan B., Hurrell E., Bennett K., Congreve M., Magnani F., Tate C.G., Weir M., and Marshall F.H. Structure of the adenosine A(2A) receptor in complex with ZM241385 and the xanthines XAC and caffeine. *Structure.***2011**; 19:1283-1293
20. Lebon G., Warne T., Edwards P.C., Bennett K., Langmead C.J., Leslie A.G., and Tate C.G. Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature.***2011**; 474:521-525
21. Rasmussen S.G., Choi H.J., Fung J.J., Pardon E., Casarosa P., Chae P.S., Devree B.T., Rosenbaum D.M., Thian F.S., Kobilka T.S., Schnapp A., Konetzki I., Sunahara R.K., Gellman S.H., Pautsch A., Steyaert J., Weis W.I., and Kobilka B.K. Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. *Nature.***2011**; 469:175-180
22. Rasmussen S.G., DeVree B.T., Zou Y., Kruse A.C., Chung K.Y., Kobilka T.S., Thian F.S., Chae P.S., Pardon E., Calinski D., Mathiesen J.M., Shah S.T., Lyons J.A., Caffrey M., Gellman S.H., Steyaert J., Skiniotis G., Weis W.I., Sunahara R.K., and Kobilka B.K. Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature.***2011**; 477:549-555
23. Shimamura T., Shiroishi M., Weyand S., Tsujimoto H., Winter G., Katritch V., Abagyan R., Cherezov V., Liu W., Han G.W., Kobayashi T., Stevens R.C., and Iwata S. Structure of the human histamine H1 receptor complex with doxepin. *Nature.***2011**; 475:65-70
24. Granier S., Manglik A., Kruse A.C., Kobilka T.S., Thian F.S., Weis W.I., and Kobilka B.K. Structure of the delta-opioid receptor bound to naltrindole. *Nature.***2012**; 485:400-404
25. Haga K., Kruse A.C., Asada H., Yurugi-Kobayashi T., Shiroishi M., Zhang C., Weis W.I., Okada T., Kobilka B.K., Haga T., and Kobayashi T. Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature.***2012**; 482:547-551
26. Hanson M.A., Roth C.B., Jo E., Griffith M.T., Scott F.L., Reinhart G., Desale H., Clemons B., Cahalan S.M., Schuerer S.C., Sanna M.G., Han G.W., Kuhn P., Rosen H., and Stevens R.C. Crystal structure of a lipid G protein-coupled receptor. *Science.***2012**; 335:851-855

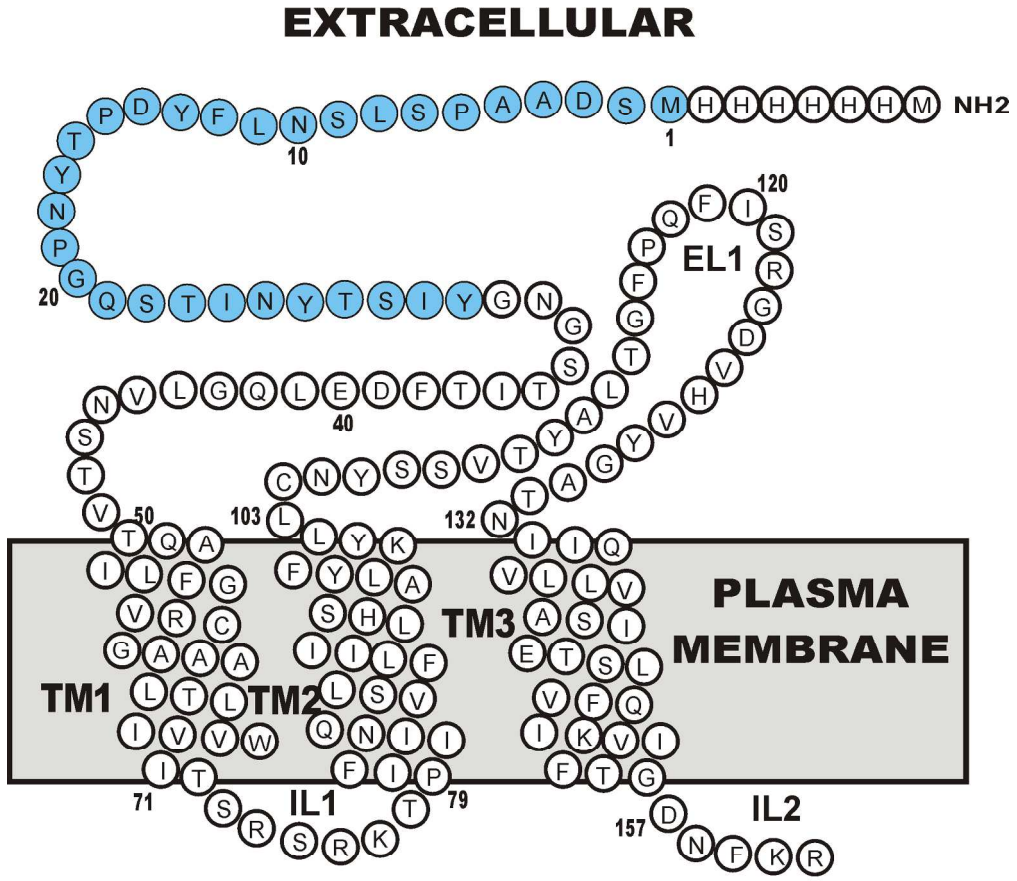
27. Kruse A.C., Hu J., Pan A.C., Arlow D.H., Rosenbaum D.M., Rosemond E., Green H.F., Liu T., Chae P.S., Dror R.O., Shaw D.E., Weis W.I., Wess J., and Kobilka B.K. Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature.***2012**; 482:552-556
28. Manglik A., Kruse A.C., Kobilka T.S., Thian F.S., Mathiesen J.M., Sunahara R.K., Pardo L., Weis W.I., Kobilka B.K., and Granier S. Crystal structure of the micro-opioid receptor bound to a morphinan antagonist. *Nature.***2012**; 485:321-326
29. Thompson A.A., Liu W., Chun E., Katritch V., Wu H., Vardy E., Huang X.P., Trapella C., Guerrini R., Calo G., Roth B.L., Cherezov V., and Stevens R.C. Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. *Nature.***2012**; 485:395-399
30. White J.F., Noinaj N., Shibata Y., Love J., Kloss B., Xu F., Gvozdenovic-Jeremic J., Shah P., Shiloach J., Tate C.G., and Grisshammer R. Structure of the agonist-bound neurotensin receptor. *Nature.***2012**; 490:508-513
31. Wu H., Wacker D., Mileni M., Katritch V., Han G.W., Vardy E., Liu W., Thompson A.A., Huang X.P., Carroll F.I., Mascarella S.W., Westkaemper R.B., Mosier P.D., Roth B.L., Cherezov V., and Stevens R.C. Structure of the human kappa-opioid receptor in complex with JDTic. *Nature.***2012**; 485:327-332
32. Zhang C., Srinivasan Y., Arlow D.H., Fung J.J., Palmer D., Zheng Y., Green H.F., Pandey A., Dror R.O., Shaw D.E., Weis W.I., Coughlin S.R., and Kobilka B.K. High-resolution crystal structure of human protease-activated receptor 1. *Nature.***2012**; 492:387-392
33. Huang J., Chen S., Zhang J.J., and Huang X.Y. Crystal structure of oligomeric beta1-adrenergic G protein-coupled receptors in ligand-free basal state. *Nat Struct Mol Biol.***2013**; 20:419-425
34. Hollenstein K., Kean J., Bortolato A., Cheng R.K., Dore A.S., Jazayeri A., Cooke R.M., Weir M., and Marshall F.H. Structure of class B GPCR corticotropin-releasing factor receptor 1. *Nature.***2013**; 499:438-443
35. Siu F.Y., He M., de Graaf C., Han G.W., Yang D., Zhang Z., Zhou C., Xu Q., Wacker D., Joseph J.S., Liu W., Lau J., Cherezov V., Katritch V., Wang M.W., and Stevens R.C. Structure of the human glucagon class B G-protein-coupled receptor. *Nature.***2013**; 499:444-449
36. Wang C., Wu H., Katritch V., Han G.W., Huang X.P., Liu W., Siu F.Y., Roth B.L., Cherezov V., and Stevens R.C. Structure of the human smoothened receptor bound to an antitumour agent. *Nature.***2013**; 497:338-343
37. Dore A.S., Okrasa K., Patel J.C., Serrano-Vega M., Bennett K., Cooke R.M., Errey J.C., Jazayeri A., Khan S., Tehan B., Weir M., Wiggin G.R., and Marshall F.H. Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. *Nature.***2014**; 511:557-562
38. Wu H., Wang C., Gregory K.J., Han G.W., Cho H.P., Xia Y., Niswender C.M., Katritch V., Meiler J., Cherezov V., Conn P.J., and Stevens R.C. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science.***2014**; 344:58-64
39. Zhang J., Zhang K., Gao Z.G., Paoletta S., Zhang D., Han G.W., Li T., Ma L., Zhang W., Muller C.E., Yang H., Jiang H., Cherezov V., Katritch V., Jacobson

K.A., Stevens R.C., Wu B., and Zhao Q. Agonist-bound structure of the human P2Y12 receptor. *Nature.***2014**; 509:119-122

40. Zhang K., Zhang J., Gao Z.G., Zhang D., Zhu L., Han G.W., Moss S.M., Paoletta S., Kiselev E., Lu W., Fenalti G., Zhang W., Muller C.E., Yang H., Jiang H., Cherezov V., Katritch V., Jacobson K.A., Stevens R.C., Wu B., and Zhao Q. Structure of the human P2Y12 receptor in complex with an antithrombotic drug. *Nature.***2014**; 509:115-118
41. Tian C., Breyer R.M., Kim H.J., Karra M.D., Friedman D.B., Karpay A., and Sanders C.R. Solution NMR spectroscopy of the human vasopressin V2 receptor, a G protein-coupled receptor. *J Am Chem Soc.***2005**; 127:8010-8011
42. Tian C., Breyer R.M., Kim H.J., Karra M.D., Friedman D.B., Karpay A., and Sanders C.R. Solution NMR spectroscopy of the human vasopressin V2 receptor, a G protein-coupled receptor. *J Am Chem Soc.***2006**; 128:5300
43. Park S.H., Das B.B., Casagrande F., Tian Y., Nothnagel H.J., Chu M., Kiefer H., Maier K., De Angelis A.A., Marassi F.M., and Opella S.J. Structure of the chemokine receptor CXCR1 in phospholipid bilayers. *Nature.***2012**; 491:779-783
44. Etzkorn M., Martell S., Andronesi O.C., Seidel K., Engelhard M., and Baldus M. Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy. *Angew Chem Int Ed Engl.***2007**; 46:459-462
45. Gautier A., Kirkpatrick J.P., and Nietlispach D. Solution-state NMR spectroscopy of a seven-helix transmembrane protein receptor: backbone assignment, secondary structure, and dynamics. *Angew Chem Int Ed Engl.***2008**; 47:7297-7300
46. Reckel S., Gottstein D., Stehle J., Lohr F., Verhoefen M.K., Takeda M., Silvers R., Kainosho M., Glaubitz C., Wachtveitl J., Bernhard F., Schwalbe H., Guntert P., and Dotsch V. Solution NMR structure of proteorhodopsin. *Angew Chem Int Ed Engl.***2011**; 50:11942-11946
47. Nygaard R., Zou Y., Dror R.O., Mildorf T.J., Arlow D.H., Manglik A., Pan A.C., Liu C.W., Fung J.J., Bokoch M.P., Thian F.S., Kobilka T.S., Shaw D.E., Mueller L., Prosser R.S., and Kobilka B.K. The dynamic process of beta(2)-adrenergic receptor activation. *Cell.***2013**; 152:532-542
48. Kimura T., Vukoti K., Lynch D.L., Hurst D.P., Grossfield A., Pitman M.C., Reggio P.H., Yeliseev A.A., and Gawrisch K. Global fold of human cannabinoid type 2 receptor probed by solid-state 13C-, 15N-MAS NMR and molecular dynamics simulations. *Proteins.***2014**; 82:452-465
49. Gautier A., Mott H.R., Bostock M.J., Kirkpatrick J.P., and Nietlispach D. Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy. *Nat Struct Mol Biol.***2010**; 17:768-774
50. Neumoin A., Cohen L.S., Arshava B., Tantry S., Becker J.M., Zerbe O., and Naider F. Structure of a double transmembrane fragment of a G-protein-coupled receptor in micelles. *Biophys J.***2009**; 96:3187-3196
51. Caroccia K.E., Estephan R., Cohen L.S., Arshava B., Hauser M., Zerbe O., Becker J.M., and Naider F. Expression and biophysical analysis of a triple-transmembrane domain-containing fragment from a yeast G protein-coupled receptor. *Biopolymers.***2011**; 96:757-771

52. Bax A., Clore G.M., and Gronenborn A.M. H-1-H-1 correlation via isotropic mixing of C-13 magnetization, a new 3-dimensional approach for assigning H-1 and C-13 spectra of C-13-enriched proteins. *J Magn Reson.***1990**; 88:425-431
53. Kay L.E., Xu G., Singer A., Muhandiram D., and Formankay J. A gradient-enhanced HCCH-TOCSY experiment for recording side-chain 1H and 13C correlations in H2O samples of proteins. *J Magn Reson B.***1993**; 101:333-337
54. Montelione G.T., Lyons B.A., Emerson S.D., and Tashiro M. An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically-enriched proteins. *J Am Chem Soc.***1992**; 114:10974-10975
55. Tugarinov V., and Kay L.E. Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *Journal of the American Chemical Society.***2003**; 125:13868-13878
56. Johnson B.A., and Blevins R.A. NMR View: A computer program for the visualization and analysis of NMR data. *J Biomol NMR.***1994**; 4:603-614
57. Keller R. 2004. The Computer Aided Resonance Assignment. C. Verlag, editor, Goldau, Switzerland.
58. Tugarinov V., Hwang P.M., and Kay L.E. Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. *Annu Rev Biochem.***2004**; 73:107-146
59. Shen Y., Delaglio F., Cornilescu G., and Bax A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR.***2009**; 44:213-223
60. Herrmann T., Guntert P., and Wuthrich K. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J Biomol NMR.***2002**; 24:171-189
61. Herrmann T., Guntert P., and Wuthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol.***2002**; 319:209-227
62. Guntert P. Automated NMR structure calculation with CYANA. *Methods Mol Biol* .**2004**; 278:353-378
63. Cohen L.S., Arshava B., Estephan R., Englander J., Kim H., Hauser M., Zerbe O., Ceruso M., Becker J.M., and Naider F. Expression and biophysical analysis of two double-transmembrane domain-containing fragments from a yeast G protein-coupled receptor. *Biopolymers.***2008**; 90:117-130
64. Bohm G., Muhr R., and Jaenicke R. Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng.***1992**; 5:191-195
65. Noggle R.C., and Schrimer R.E. *The Nuclear Overhauser Effect: Chemical Applications*. Adacemic Press, New York, NY, 1971.
66. Hofmann K., and Stoffel W. TMbase-A database of membrane spanning protein segments. *Biol Chem.***1993**; 374:
67. Hirokawa T., Boon-Chieng S., and Mitaku S. SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics.***1998**; 14:378-379

68. Krogh A., Larsson B., von Heijne G., and Sonnhammer E.L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.***2001**; 305:567-580
69. Juretic D., Zoranic L., and Zucic D. Basic charge clusters and predictions of membrane protein topology. *J Chem Inf Comput Sci.***2002**; 42:620-632
70. Eilers M., Hornak V., Smith S.O., and Konopka J.B. Comparison of class A and D G protein-coupled receptors: common features in structure and activation. *Biochemistry.***2005**; 44:8959-8975
71. Shi C., Kaminskyj S., Caldwell S., and Loewen M.C. A role for a complex between activated G protein-coupled receptors in yeast cellular mating. *Proc Natl Acad Sci, U S A.***2007**; 104:5395-5400
72. Shi C., Kendall S.C., Grote E., Kaminskyj S., and Loewen M.C. N-terminal residues of the yeast pheromone receptor, Ste2p, mediate mating events independently of G1-arrest signaling. *J Cell Biochem.***2009**; 107:630-638
73. Uddin M.S., Kim H., Deyo A., Naider F., and Becker J.M. Identification of residues involved in homodimer formation located within a beta-strand region of the N-terminus of a Yeast G protein-coupled receptor. *J Recept Signal Transduct Res.***2012**; 32:65-75
74. Chou P.Y., and Fasman G.D. Prediction of protein conformation. *Biochemistry.***1974**; 13:222-245
75. Chou P.Y., and Fasman G.D. Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins. *Biochemistry.***1974**; 13:211-222
76. Cole C., Barber J.D., and Barton G.J. The Jpred 3 secondary structure prediction server. *Nucleic Acids Res.***2008**; 36:W197-201
77. Zerbe O. First solution structures of seven-transmembrane helical proteins. *Angew Chem Int Ed Engl.***2012**; 51:860-861
78. Gurevich V.V., and Gurevich E.V. GPCR monomers and oligomers: it takes all kinds. *Trends Neurosci.***2008**; 31:74-81
79. Fanelli F., and Felline A. Dimerization and ligand binding affect the structure network of A(2A) adenosine receptor. *Biochim Biophys Acta.***2011**; 1808:1256-1266
80. Kasai R.S., and Kusumi A. Single-molecule imaging revealed dynamic GPCR dimerization. *Curr Opin Cell Biol.***2014**; 27:78-86



Structural characterization of integral membrane proteins is an arduous task that requires patience and innovation for progress. In this paper we characterize the secondary structure of a three transmembrane domain fragment of the Ste2p G protein-coupled receptor from *Saccharomyces cerevisiae*. Methods are introduced to increase expression and stability of this protein for future structural characterization.

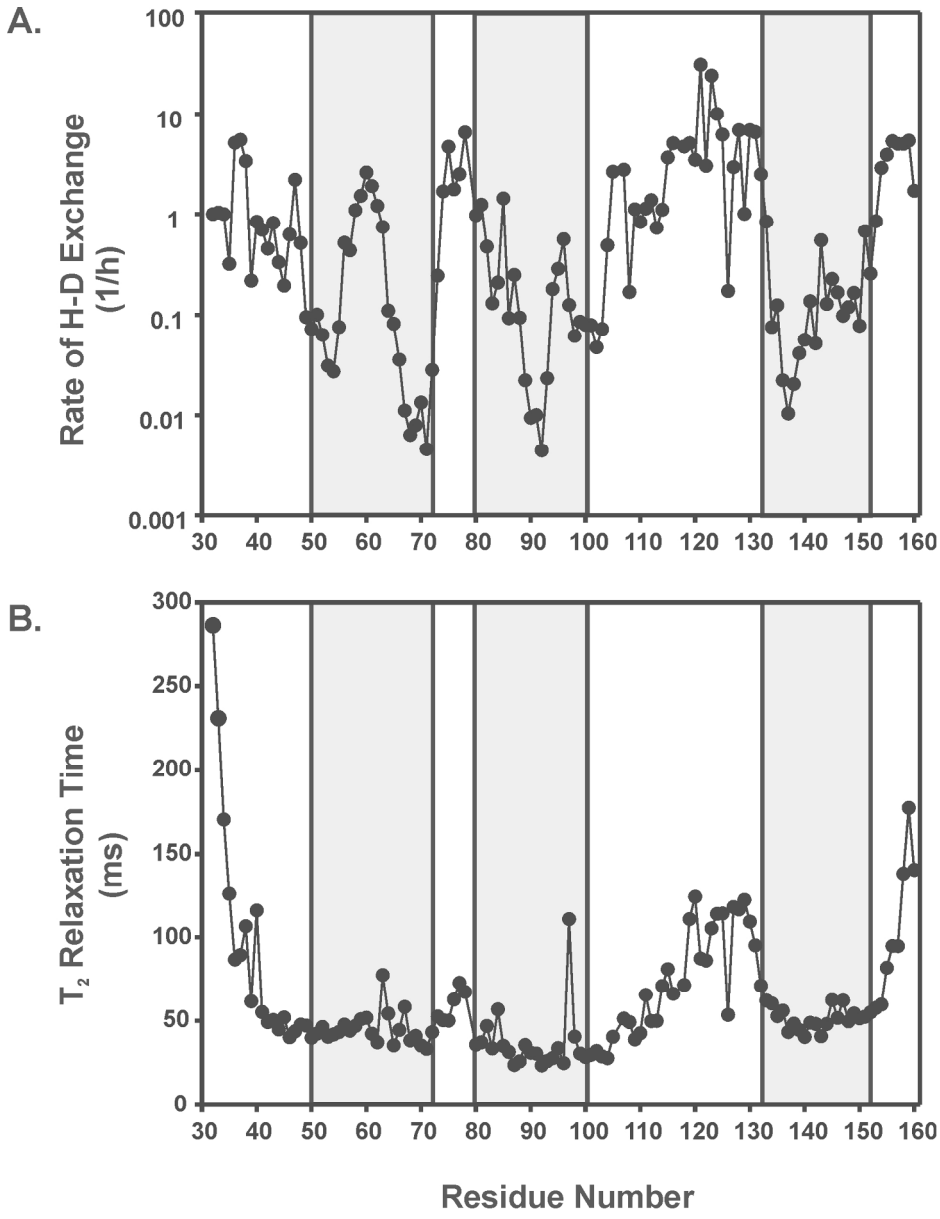


Figure 1. Evaluation of the secondary structure of Ste2p TM1-TM3 (G31-R161) by Hydrogen-Deuterium amide Exchange and T_2 Relaxation. A) Exchange times were calculated from a series of [^{15}N,^{1}H]-HSQC experiments conducted at 45°C, and plotted as a function of residue number with an inverse logarithmic scale. B) The T_2 relaxation times for all residues in TM1-TM3 were measured using the series of experiments conducted at 45°C as described. The boxed helical boundaries are those calculated by sequence-based TM prediction software (average of four programs).

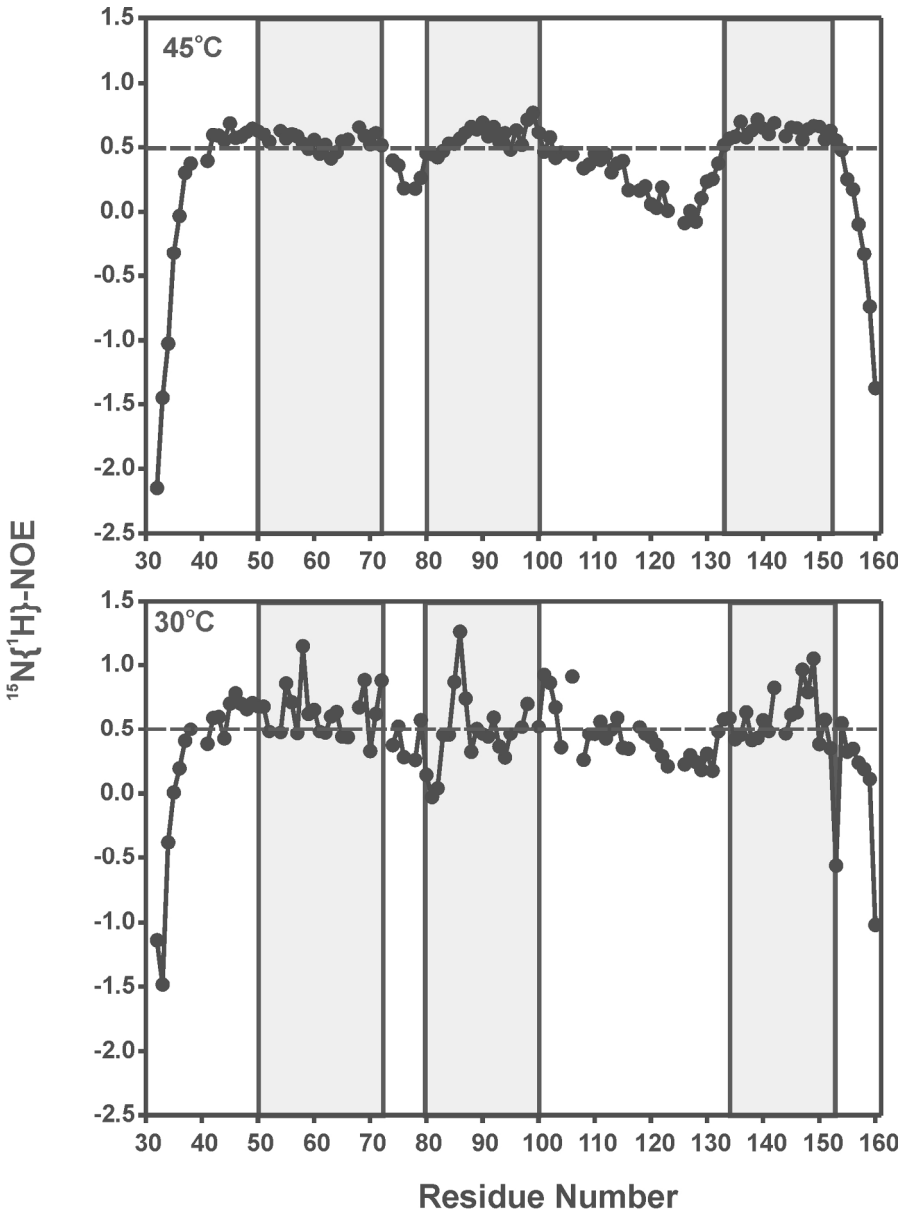


Figure 2. Evaluation of the relative mobility of the Ste2p TM1-TM3 (G31-R161) construct by ^{15}N{^{1}H}-NOE. [^{15}N,^{1}H]-HSQC experiments were conducted at 45°C and 30°C with and without amide proton irradiation. The ratio of the peak amplitude with the pulse to the amplitude without the pulse for every residue in the fragment was calculated and plotted as a function of residue number. The boxed helical boundaries are those calculated by sequence-based TM prediction software (average of four programs).

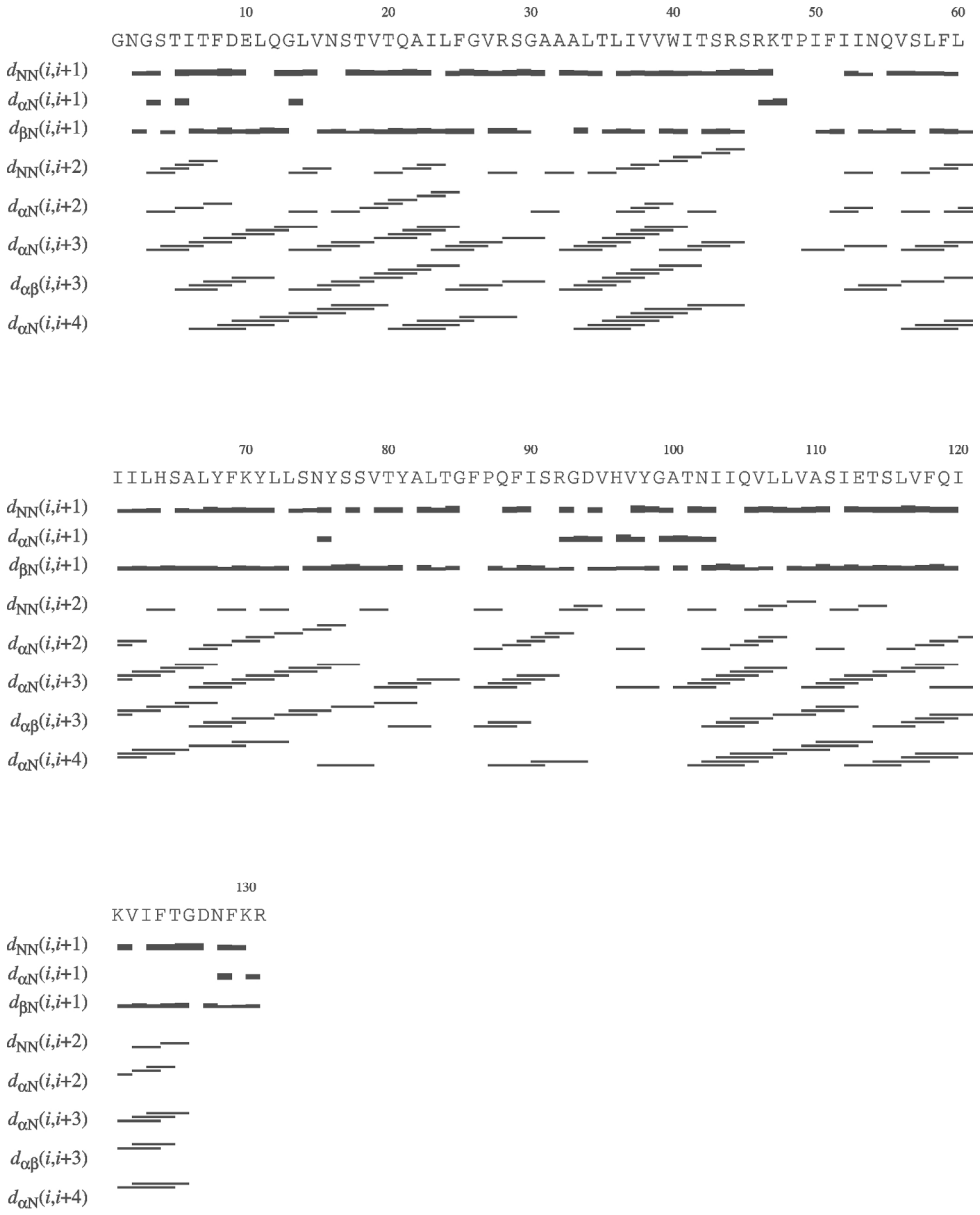


Figure 3: Inter-residue NOE connectivities for the CYANA structure calculation of Ste2p TM1-TM3 (G31-R161). The ATNOS-CANDID component of the UNIO software suite was used for automatic assignment of the [^{13}C]- and [^{15}N]-resolved NOESY experiments. Connectivities are displayed as a function of residue.

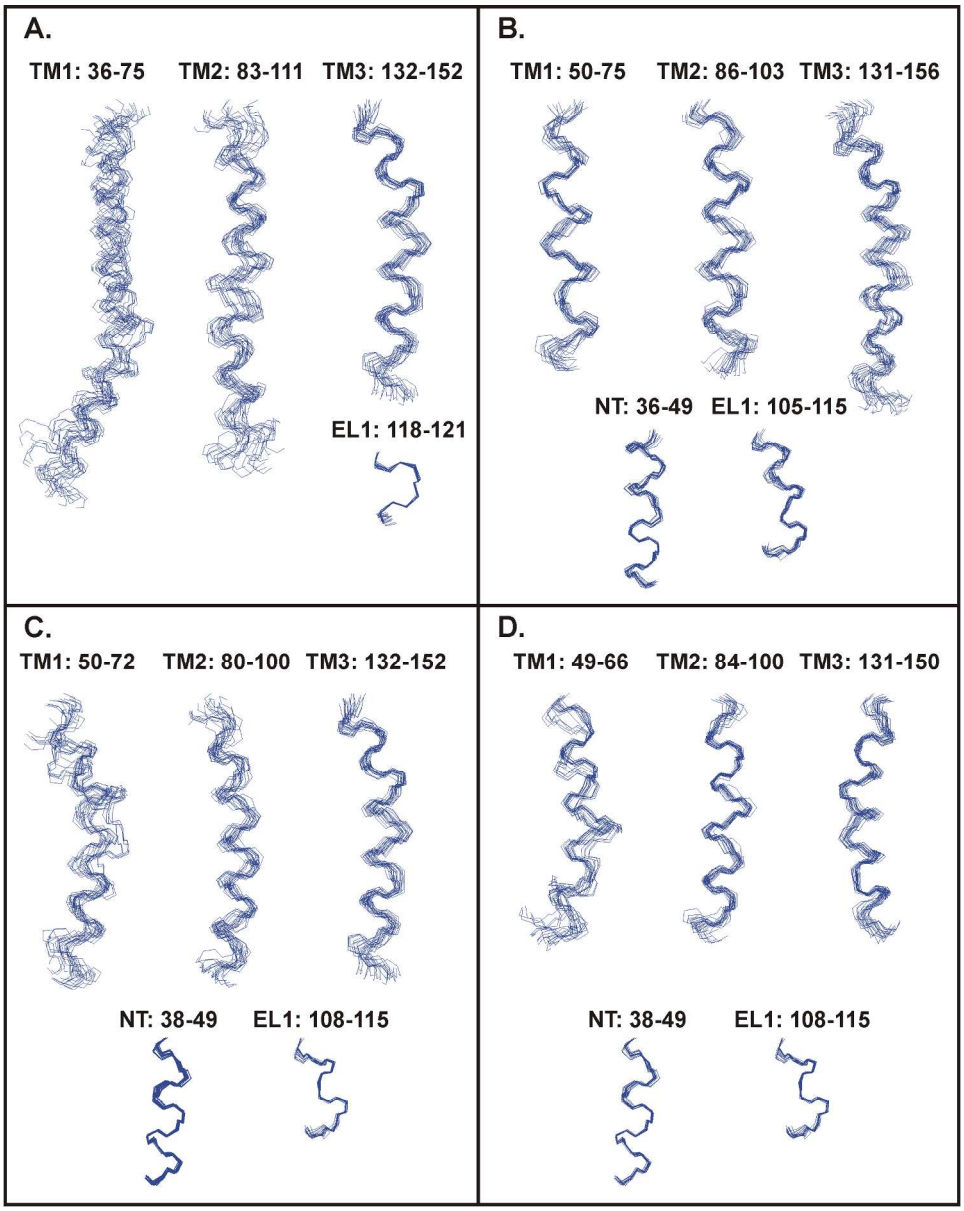


Figure 4: Convergence of the lowest 20 energy CYANA calculated structures for Ste2p TM1-TM3 (G31-R161). The helices were produced by fitting the helical boundaries calculated from the structure (A), those determined from inter-residue NOE connectivities (B), the calculated TM boundaries based on prediction software and the secondary shift boundaries for the NT and EL1 helices (C), and the template predicted TM boundaries and the secondary shift boundaries for the NT and EL1 helices (D).

Table 1: Calculated RMSD Values for secondary structure elements of TM1-TM3

Method	Putative Ste2p Region	Included Residues	Backbone RMSD	Heavy Atom RMSD
Structure[a]	NT-TM1	36-75	3.02 ± 1.41	3.81 ± 1.52
	TM2-EL1	83-111	2.09 ± 0.81	2.98 ± 0.84
	EL1	118-121	0.24 ± 0.11	1.17 ± 0.31
	TM3	132-152	0.99 ± 0.29	1.64 ± 0.29
NOE[b]	TM1	50-75	2.36 ± 1.16	3.05 ± 1.13
	TM2	86-103	1.03 ± 0.41	1.98 ± 0.46
	TM3	131-156	1.49 ± 0.51	2.07 ± 0.44
	NT	36-49	0.43 ± 0.18	1.31 ± 0.21
	EL1	105-115	0.69 ± 0.28	1.53 ± 0.40
Sequence-based Calculation[c]	TM1	50-72	2.10 ± 1.09	2.77 ± 1.04
	TM2	80-100	1.47 ± 0.5	2.5 ± 0.56
	TM3	132-152	0.99 ± 0.31	1.64 ± 0.29
	NT	38-49	0.39 ± 0.18	1.33 ± 0.23
	EL1	108-115	0.43 ± 0.22	1.14 ± 0.41
Eilers[d]	TM1	49-66	1.54 ± 0.79	2.1 ± 0.71
	TM2	84-100	0.82 ± 0.31	1.8 ± 0.34
	TM3	131-150	0.85 ± 0.28	1.49 ± 0.27
	NT	38-49	0.39 ± 0.18	1.33 ± 0.23
	EL1	108-115	0.43 ± 0.22	1.14 ± 0.41

[a] Boundaries determined in MOLMOL using the Kabsch-Sanders algorithm from the calculated structure.
[b]Boundaries from the calculated structure modified based on *i, i+4* NOE connectivites.
[c]Average boundaries from sequence-based TM prediction software.
[d]Boundaries from the rhodopsin-templated model of Ste2p

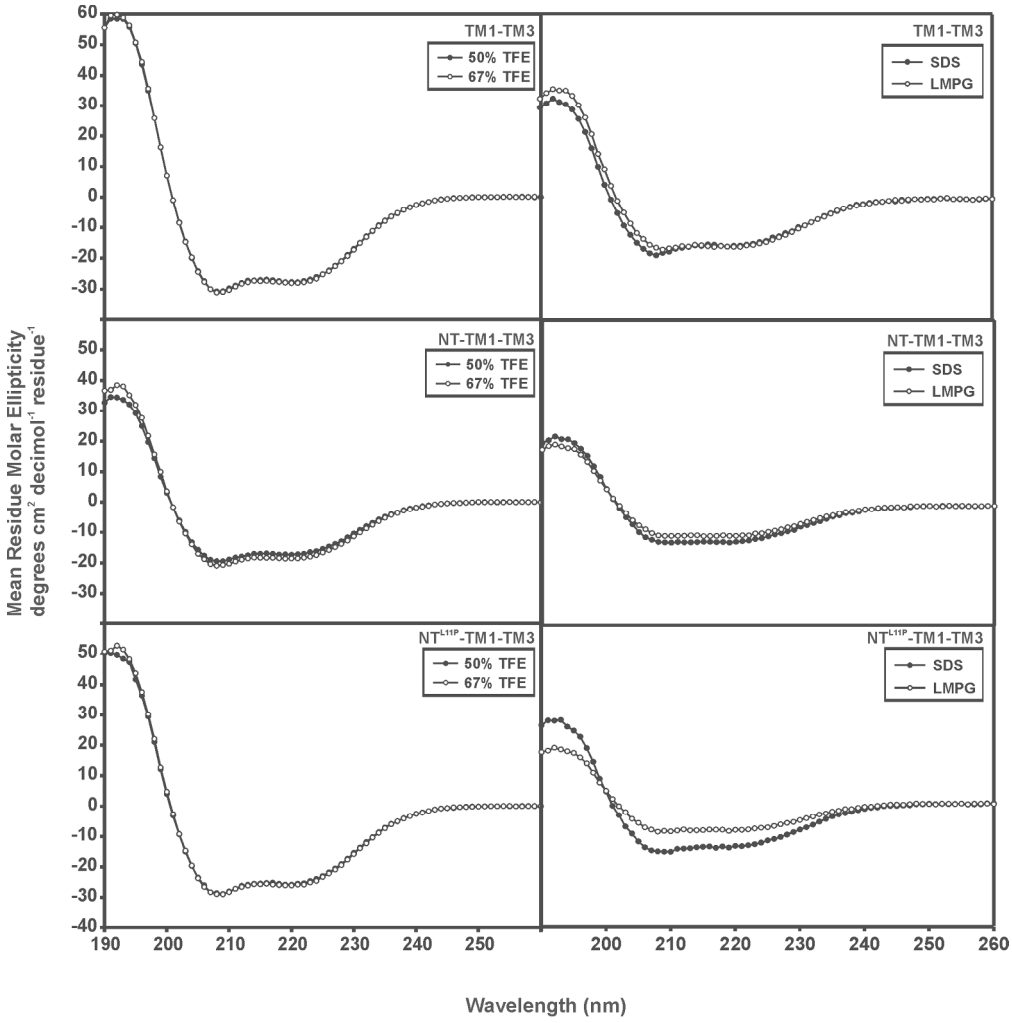


Figure 5. Secondary structure analysis of TM1-TM3, NT-TM1-TM3, and NT^{L11P}-TM1-TM3 using circular dichroism. Left panels: Lyophilized peptides were solubilized in 50% and 67% TFE as described. Right panels: Lyophilized peptides were solubilized in 20 mM SDS and LMPG as described. The presented spectra represent a solvent-subtracted average of four scans.

Table 2. Percentages of secondary structures based on deconvolution of CD data using CDNN[64]

Peptide	TM1-TM3		NT-TM1-TM3		NT^{L11P}-TM1-TM3	
	α-helix	β-sheet	α-helix	β-sheet	α-helix	β-sheet
50% TFE	86.0%	1.6%	56.2%	6.1%	80.0%	2.2%
67% TFE	86.5%	1.5%	61.1%	5.1%	81.2%	2.1%
SDS	53.6%	6.8%	38.8%	12.7%	46.7%	9.1%
LMPG	55.3%	6.4%	34.2%	15.9%	32.0%	18.3%